DOCUMENT RESUME

ED 287 881 TM 870 635

AUTHOR Neel, John H.
TITLE Sample Size for Variance in a Power Calculation.
PUB DATE Apr 87
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, DC, April 20-24, 1987).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Analysis of Variance; *Error of Measurement; *Mathematical Models; Sample Size; *Statistical Significance; Statistical Studies
IDENTIFIERS *Power (Statistics)

ABSTRACT

Determination of statistical power for analysis of variance procedures requires five elements: (1) significance level; (2) effect size; (3) number of means; (4) error variance; and (5) sample size. Significance levels are traditionally chosen to be 0.5, .01, or .001. Effect size is not discussed in this paper. The number of means is determined by the particular experimental setting. Sample size is guessed at until the desired power is obtained, leaving error variance to be established before the calculation can proceed. Establishing the goodness of the estimate of the error variance is the focus here, in an explanation of what constitutes a good estimate. Discussion is restricted to a consideration of error variances in one-way design. (RR)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED287881

# SAMPLE SIZE FOR VARIANCE IN A POWER CALCULATION

John H. Neel

Georgia State University

Presented at the American Educational Research Association Convention, Washington, D.C. April 22, 1987

TM 870635

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John H. Neel

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

## Introduction

Methods for determining statistical power and sample size are covered in many texts (Cohen 1977; Kirk,1982; Glass and Stanley, 1970). Researchers are encouraged to use these techniques in order to have adequate sample sizes for experiments. See Kirk (1982) for a for a good presentation of the techniques of estimating power. Five quantities are necessary in order to calculate power for analysis of variance procedures. These are:

1. significance level,
2. effect size,
3. number of means,
4. error variance, and
5. sample size.

The significance level is usually chosen by tradition to be .05, .01, or .001. The effect size is not the focus of this paper. Some help with effect sizes may be found in Cohen (1977). The number of means is determined by the particular experimental setting. Sample size is guessed at until the desired power is obtained. This leaves the error variance to be established before the calculation can proceed. Most texts offer little advice on how to establish the error variance. Some refer to a "good" estimate without defining what constitutes a good estimate. Establishing the goodness of the estimate of the error variance is the focus of this paper. The paper is restricted to a consideration of error variances in one-way designs.

Goodness of Variance Estimators

The error variance in an analysis of variance has a chi-square distribution (Scheffe, 1959 P. 125). As such, and given random sampling, the goodness of the estimate is dependent only on the sample size or degrees of freedom used to estimate the variance. The variance will be good in that it will be unbiased, and of minimum variance. Most introductory theory of statistics texts introduce these concepts with respect to variance. These discussions invariably assume a normal distribution and a normal distribution of the errors from which the error variance is calculated is assumed in this paper. Indeed, the standard normality, homogeneity of variance, independence, and random sampling assumptions of analysis of variance are all presumed here. Given these assumptions, the distribution of the variance estimate is well known as the chi-square distribution and thus depends only on the sample size, or equivalently, the degrees of freedom of the estimate.

The question then arises, how large must the sample size be in order to estimate the error variance for use in a power calculation. To answer this question some criterion must be established. The criteria selected here is based on the correctness of the power calculation.

What powers do experimenters wish to estimate for their experiments? The lowest power of interest seems to be .5 to the author. This is of interest because the argument is sometimes made that the experiment was not worth doing since the power was such that a coin could be tossed and be as powerful as the experiment. The highest commonly used value of alpha (type one error) is .001. If we assign the same value to beta, the probability of a type II error, then a power of .999 is the highest power of interest (power = 1 - beta = 1 - .001 =

.999). With these limits in mind, powers to examine in this paper were selected to be .50, .75, .85, .90, .95, .99, and .999.

What is a reasonable accuracy for these selected powers? Undoubtedly there could be many answers to that question. For the purposes of this paper it was decided that if p is the estimate of the desired power, P, then p should be within (1-P)/2 of P. It was further decided that the probability of p being within these limits should be .95. These are relatively arbitrary decisions. Others could easily state other criteria. The reasoning here is first that the estimation of power is at best a tenuous matter. Both the error variance and the effect size must usually be estimated. As the estimation errs the power calculation errs. Given this tenuousness of the process, these bounds on the obtained probability seemed reasonable. Table 1 lists the probabilities and the error bounds selected for this study.

Parameters selected for study

In addition to the desired powers selected in the previous section, values of alpha, effect size, and the number of means needed to be selected. The values selected for alpha were the commonly used ones mentioned above: .05, .01, .001. The values of effect size selected were the small medium and large sizes discussed by Cohen (1977, P. 24). The range of means selected was 2 to 8. This results in 441 combinations of desired power, significance level, effect size, and number of means to be studied.

METHODS

Given a fixed effect size, alpha, number of means, and desired power; the resulting power is a function of the error variance. That is, the value found for power will vary in a one-to-one correspondence with the error variance estimate. The probability distribution of the power is thus the induced probability distribution from the distribution of the error variance (See Meyer, 1965, Pp. 50,72) for a discussion of induced probabilities).

TABLE 1

SELECTED POWER AND BOUNDS TO POWER

| Power | Bound |
|---|---|
| .50 | .250 - .750 |
| .75 | .625 - .875 |
| .85 | .775 - .925 |
| .90 | .850 - .950 |
| .95 | .925 - .975 |
| .99 | .985 - .995 |
| .999 | .9985 - .9995 |

The reader may find it strange to think of power as a random variable. We are used to thinking of power as a value to be calculated or read from a table. This it usually is, but there is the assumption when we calculate or read the table that all the values we use in our calculation or to the read the table are correct. When one of the values we use is a random variable, then power itself becomes a random variable and the distribution of power thus depends on the distribution of the random variable. For the current study, it is assumed that all values ( alpha, effect size, number of means, and desired power) are known constants. Since the error variance has been estimated, it is the only random variable and the distribution of power is thus dependent on the distribution of the error variance.

To find the distribution of power, it is only necessary to find the distribution of the error variance. Tables of the distribution of power were prepared using the values from the chi-square distribution. Since the small, medium, and large effect sizes of Cohen (1977, P. 24) are in effect scale free, the population value of the error variance was assumed to be 1.0 without any loss of generality.

Then the formula for chi-square was solved to yield

$$s^2 = (\sigma^2 \chi^2)/(n-1) \ .$$

Using a chi-square table, the .025 and .975 values could be substituted into the above equation together with n to yield the values of $s^2$ which bound a probability of .95. The procedure then was to guess n's and see what power resulted. The smallest and largest n's which would yield power just within ( P-(1-p)/2, P+(1-P)/2 ) were

,respectivaly, the n's for the .025 and .975 values of chi-square and thus , due to the induction, for power. The n's for the .025 and .975 points differ and to be conservative, one takes the larger value n.

If $s^2$ were based on n-1 degrees of freedom, the estimate of P would thus have probability .95 of being within the bound shown in Table 1. These calculations were performed for all combinations of desired power = .50, .75, .85, .90, .95, .99, .999; effect size = .20, .50, .80; alpha= .05, .01, .001; and number of means = 2, 3, 4, 5, 6, 7, 8. This results in 441 values which were tabled.

A computer program was written in C under a UNIX operating system on a TANDY Model 6000 microcomputer to do the calculations. Biometrika (Pearson and Hartley, 1970) tables were used up to degrees of freedom 100 for values of chi-square. Then the normal approximation to the chi-square was used (Kirk, 1982, P. 821).

A routine by Woodard and Overall (1976) was used to do the power calculations. The routine was rewritten from FORTRAN to C and modified to achieve greater speed. The modification did not effect the calculations. Results between the modified and unmodified versions agreed to 6 digits. The routine requires values of the F distribution. These were also taken from the Biometrika (Pearson and Hartley, 1970) tables. Tabled values or interpolated values were used up to values of denominator degrees of freedom of 120. Above 120 the value for infinite denominator degrees of freedom was used. ( When a smaller table was used it produced some "jumps" in the power calculation so the larger was necessary.)

## Accuracy

The routine by Woodard and (1976) involves 2 approximations and the values of chi-square above 100 degrees of freedom involve an approximation. With so much approximation a legitimate question is just how accurate are the calculations. A number of values (~20) were calculated by hand using the power charts in Kirk (1982, Pp. 832-839). These calculations agreed to within the ability to read the charts.

## Results

Figure 1 was constructed from a table of the necessary degrees of freedom for obtaining the desired power of .05, alpha of .05, effect size of .25, and two means. Figure 2 used the same parameters except for desired power of .75 and Figure 3 gives the same data for a desired power of .95. The curved lines on each Figure represent .005, .001,.01, .025, .050, .100, .250, .500, .750, .90, .95, .975, .99, .995, and .999, respectively, reading from bottom to top. The horizontal line is drawn through the desired power.

One reads the Figures as follows. On Figure 1 let the degrees of freedom for the error variance ( v for MSE ) be 10. Reading vertically to the .50 line and then left to the obtained power, we find obtained power of about .48. In other words, for the given setting the median obtained power is .48 when the desired power is .50. (The median obtained power being less than the desired power reflects the general nonsymetry of the F and noncentral F distributions.) Similarly, reading at degrees of

FIGURE 1

PROBABILITY BOUNDS FOR OBTAINED POWER
GIVEN DESIRED POWER OF 0.500, $\upsilon 1 = 1$,
$\alpha = 0.05$, AND EFFECT SIZE=0.25

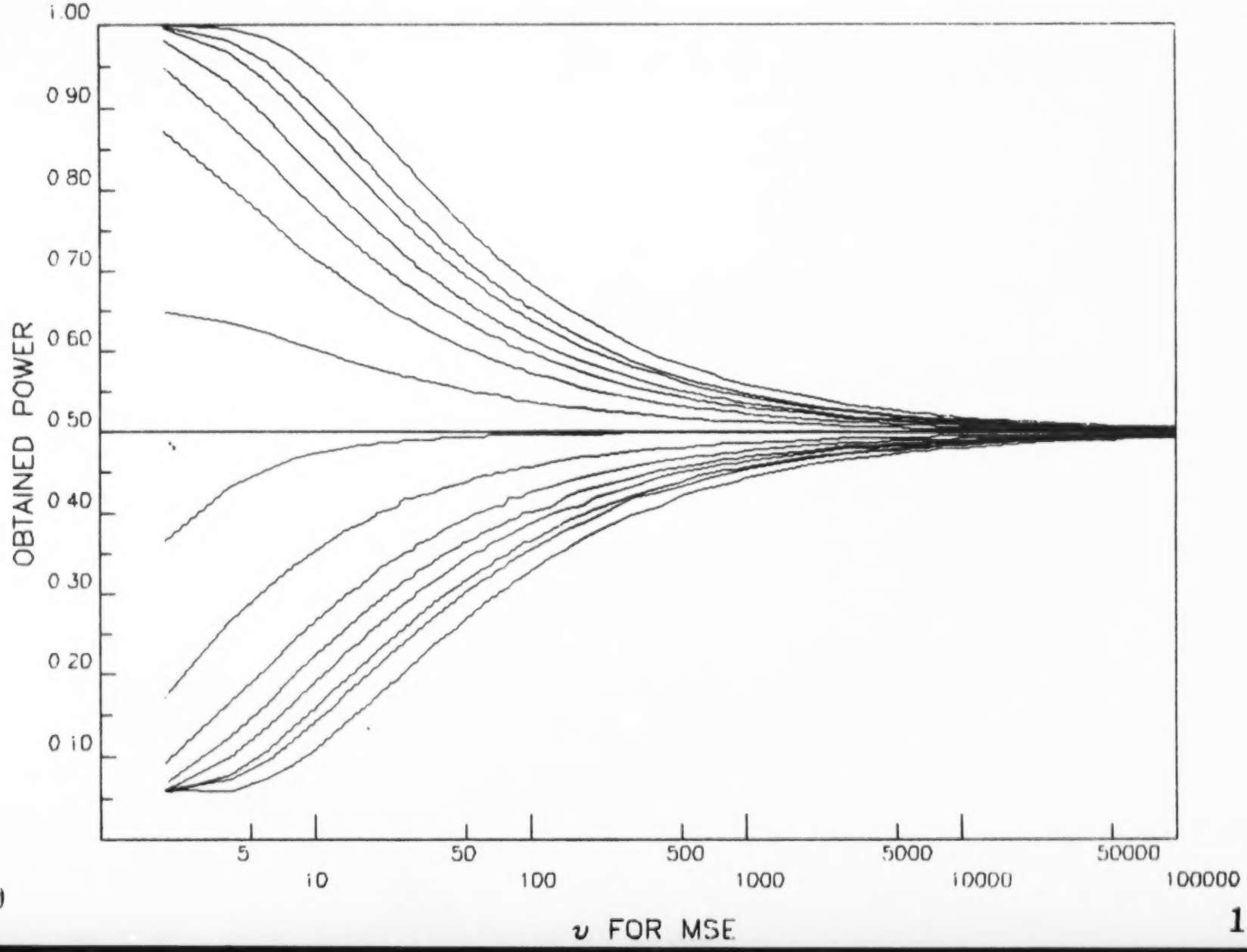


FIGURE 2

PROBABILITY BOUNDS FOR OBTAINED POWER GIVEN DESIRED POWER OF 0.750, $\upsilon 1 = 1$, $\alpha = 0.05$, AND EFFECT SIZE=0.25

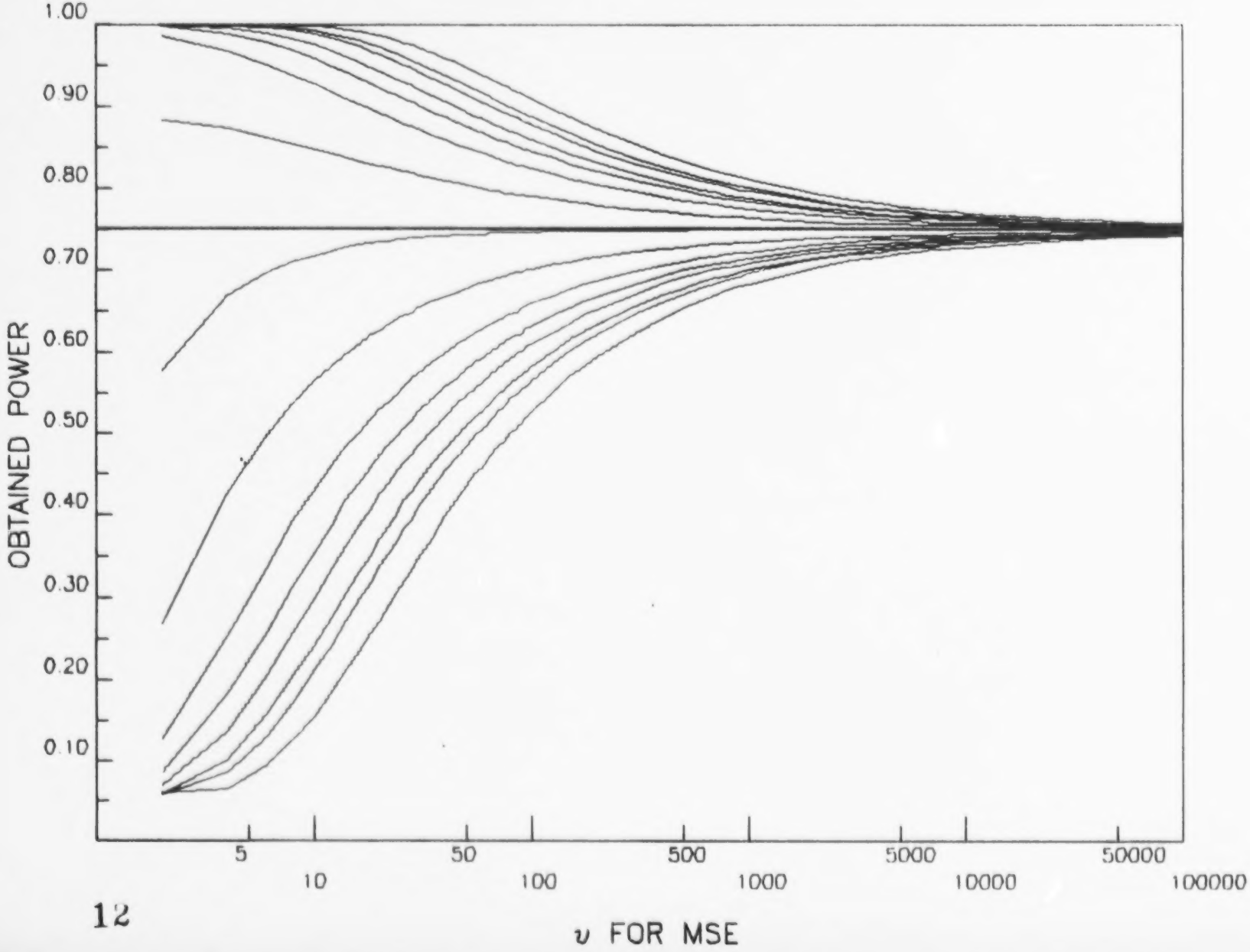
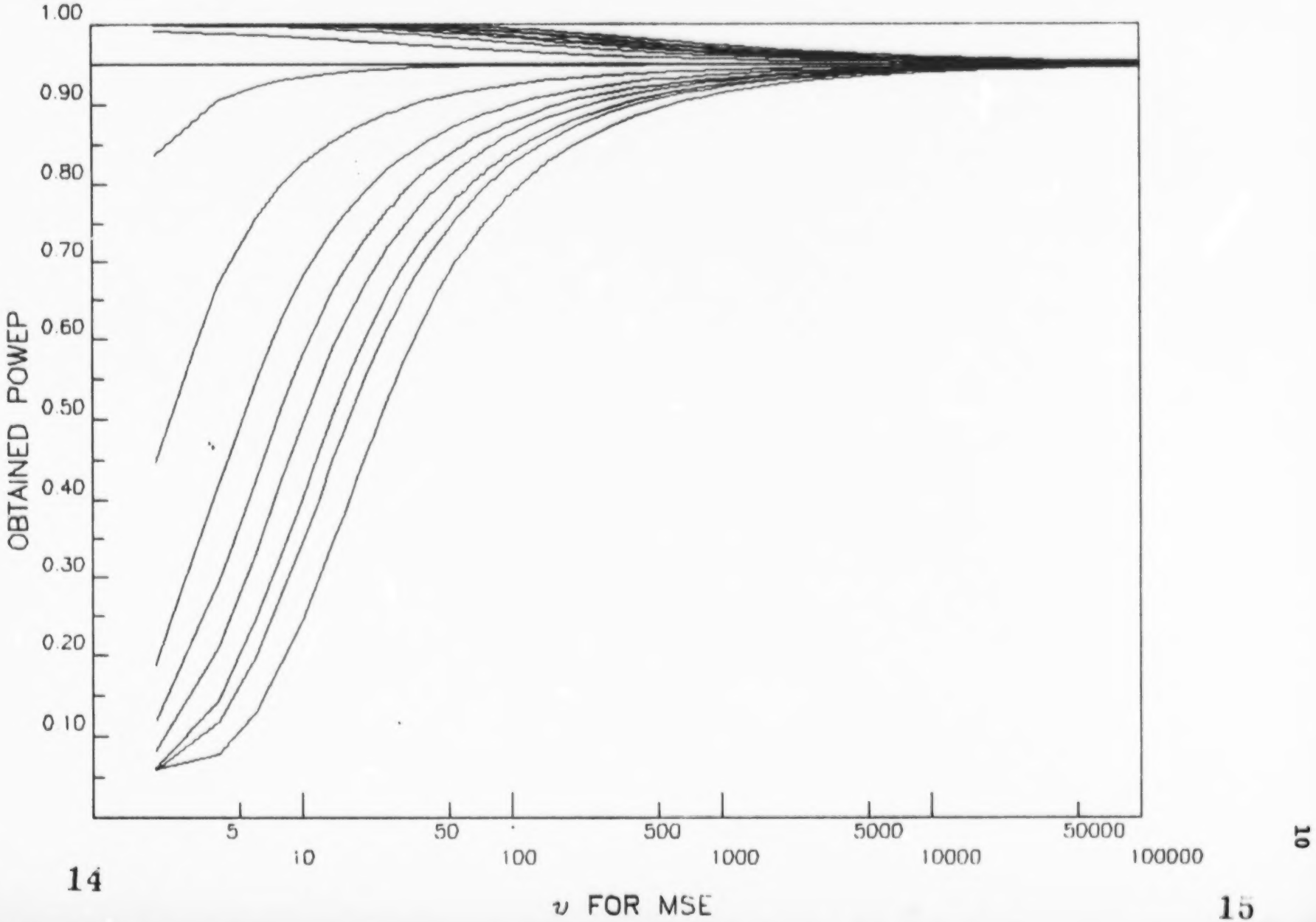


FIGURE 3

PROBABILITY BOUNDS FOR OBTAINED POWER GIVEN DESIRED POWER OF 0.950, $\upsilon 1 = 1$, $\alpha = 0.05$, AND EFFECT SIZE=0.25



freedom 100 and up to the 4th and 12th curved lines we find that obtained power will be, roughly, within .38 and .63 with probability .95.

Figures 1 through 3 are at best difficult to read. Moreover, how shall they be interpreted? What probability bound shall be used and what accuracy of estimated power? It was to answer these questions that the previously discussed limits of (1-P)/2 and probability .95 were selected. Using these limits, it is possible to take two points from each of the Figures. The maximum of these two points is the required power. By establishing these criteria, the 441 Figures at all combinations of desired power, alpha, effect size, and number of means could be reduced to 441 numbers. The resulting degrees of freedom for estimating the error variance are tabled in TABLES 1, 2, and 3 for alpha of .05, .01, and .001 respectively. Several trends are evident in the Tables. First, the required degrees of freedom rises as the desired power rises. This is not surprising since the acceptable error declines as the desired power rises. Second, the required degrees of freedom rises as alpha rises. Third, there appears to be a modest rise in the required degrees of as the number of means rises. Fourth, there is a slight rise in the necessary degrees of freedom as the effect size rises.

Use of the Tables

Use of the tables is simple. One selects alpha, suspected effect size, number of means, and desired power. Using these values the sample size which would yield the desired power within plus or minus half the distance from that power to 1 can be read from the tables. One can thus decide how good an estimator to use to estimate the error variance.

## Discussion

The results of this study are not particularly encouraging for researchers. When determining sample size for power of .5 the required degrees of freedom of the error variance range from 19 to 73. However, when the desired power is .95 the required degrees of freedom range from 417 to 1856. This would say that the practice of running a pilot study with a few dozen subjects and using the resulting error variance to estimate power is probably not advisable in that the estimate would not meet the requirements given here.

Whether the requirements here are too stringent remains to be seen. A preliminary calculation seeking a .75 probability bound of the same size on the desired powers is not reported here but also gave large requirements.

In any case, there is at least a caution to the use of small sample sizes in estimating error variancs for a power calculation. It should be noted that this difficulty may be avoided completely by using Cohen's (1977) definition of small, medium, and large effect sizes. It is not always satisfactory to use the Cohen's procedures. Sometimes it is desirable to work in the metric of the dependent variable. When this is the case the tables presented here offer some guidelines for choosing a sample size to estimate the error variance.

TABLE 2

DEGREES OF FREEDOM NECESSARY TO ESTIMATE ERROR VARIANCE SUCH THAT THE DESIRED POWER, P, IS ACCURATE TO WITHIN (1-P)/2 WITH PROBABILITY .05 FOR VARIOUS EFFECTS AND NUMBERS OF MEANS IN A ONE-WAY ANALYSIS OF VARIANCE.

FOR ALPHA = .05

| | | DESIRED POWER | | | | | | |
|---|---|---|---|---|---|---|---|---|
| EFFECT | # MEANS | .50 | .75 | .85 | .90 | .95 | .99 | .999 |
| .20 | | | | | | | | |
| | 2 | 20 | 128 | 267 | 417 | 761 | 1763 | 3151 |
| | 3 | 22 | 142 | 294 | 461 | 824 | 1923 | 3464 |
| | 4 | 24 | 148 | 314 | 486 | 876 | 2070 | 3669 |
| | 5 | 25 | 155 | 324 | 511 | 916 | 2138 | 3860 |
| | 6 | 26 | 160 | 333 | 515 | 926 | 2197 | 3885 |
| | 7 | 27 | 162 | 344 | 537 | 947 | 2250 | 3979 |
| | 8 | 27 | 166 | 350 | 538 | 969 | 2294 | 4071 |
| .50 | | | | | | | | |
| | 2 | 20 | 125 | 278 | 405 | 816 | 1613 | 2878 |
| | 3 | 21 | 133 | 280 | 489 | 815 | 1797 | 3429 |
| | 4 | 23 | 146 | 304 | 497 | 829 | 1909 | 3962 |
| | 5 | 24 | 160 | 308 | 491 | 865 | 1963 | 3539 |
| | 6 | 25 | 157 | 323 | 518 | 864 | 2086 | 3668 |
| | 7 | 26 | 158 | 351 | 506 | 934 | 2132 | 3826 |
| | 8 | 27 | 169 | 334 | 558 | 1022 | 2324 | 3835 |
| .80 | | | | | | | | |
| | 2 | 19 | 141 | 274 | 351 | 673 | 1527 | 2539 |
| | 3 | 23 | 133 | 346 | 506 | 700 | 1702 | 4026 |
| | 4 | 21 | 134 | 223 | 428 | 812 | 2183 | 3476 |
| | 5 | 24 | 163 | 351 | 508 | 760 | 1814 | 3733 |
| | 6 | 27 | 161 | 290 | 512 | 915 | 2224 | 4148 |
| | 7 | 26 | 164 | 342 | 547 | 823 | 2015 | 3569 |
| | 8 | 27 | 178 | 330 | 480 | 867 | 2290 | 4081 |

TABLE 3

DEGREES OF FREEDOM NECESSARY TO ESTIMATE ERROR VARIANCE SUCH THAT THE DESIRED POWER, P, IS ACCURATE TO WITHIN (1-P)/2 WITH PROBABILITY .05 FOR VARIOUS EFFECTS AND NUMBERS OF MEANS IN A ONE-WAY ANALYSIS OF VARIANCE.

FOR ALPHA = .01

| | | DESIRED POWER | | | | | | |
|---|---|---|---|---|---|---|---|---|
| EFFECT | # MEANS | .50 | .75 | .85 | .90 | .95 | .99 | .999 |
| .20 | | | | | | | | |
| | 2 | 33 | 192 | 384 | 580 | 1026 | 2299 | 4067 |
| | 3 | 38 | 213 | 423 | 638 | 1111 | 2521 | 4342 |
| | 4 | 40 | 225 | 448 | 685 | 1187 | 2642 | 4518 |
| | 5 | 42 | 232 | 467 | 699 | 1235 | 2699 | 4680 |
| | 6 | 44 | 241 | 476 | 728 | 1268 | 2762 | 4857 |
| | 7 | 45 | 248 | 491 | 735 | 1291 | 2878 | 4955 |
| | 8 | 46 | 252 | 496 | 757 | 1303 | 2893 | 5004 |
| .50 | | | | | | | | |
| | 2 | 33 | 187 | 364 | 545 | 1030 | 2256 | 4080 |
| | 3 | 38 | 215 | 425 | 648 | 1082 | 2297 | 4007 |
| | 4 | 40 | 221 | 434 | 632 | 1221 | 2800 | 4612 |
| | 5 | 42 | 240 | 441 | 716 | 1281 | 2515 | 4440 |
| | 6 | 43 | 237 | 462 | 752 | 1282 | 2689 | 4680 |
| | 7 | 45 | 236 | 493 | 733 | 1213 | 2816 | 5136 |
| | 8 | 46 | 259 | 488 | 743 | 1368 | 2763 | 5393 |
| .80 | | | | | | | | |
| | 2 | 33 | 213 | 445 | 548 | 1263 | 2155 | 3144 |
| | 3 | 40 | 192 | 353 | 599 | 1233 | 2799 | 3470 |
| | 4 | 40 | 237 | 494 | 578 | 1267 | 2421 | 5432 |
| | 5 | 38 | 226 | 513 | 726 | 1299 | 2360 | 4414 |
| | 6 | 42 | 239 | 463 | 788 | 1183 | 3142 | 5461 |
| | 7 | 42 | 220 | 479 | 740 | 1317 | 2547 | 4339 |
| | 8 | 43 | 256 | 508 | 726 | 1126 | 3254 | 5613 |

TABLE 4

DEGREES OF FREEDOM NECESSARY TO ESTIMATE ERROR VARIANCE SUCH THAT THE DESIRED POWER, P, IS ACCURATE TO WITHIN (1-P)/2 WITH PROBABILITY .05 FOR VARIOUS EFFECTS AND NUMBERS OF MEANS IN A ONE-WAY ANALYSIS OF VARIANCE.

FOR ALPHA = .001

| | | DESIRED POWER | | | | | | |
|---|---|---|---|---|---|---|---|---|
| EFFECT | # MEANS | .50 | .75 | .85 | .90 | .95 | .99 | .999 |
| .20 | | | | | | | | |
| | 2 | 53 | 279 | 537 | 808 | 1360 | 2934 | 5081 |
| | 3 | 59 | 303 | 589 | 883 | 1481 | 3205 | 5325 |
| | 4 | 62 | 321 | 621 | 912 | 1568 | 3352 | 5605 |
| | 5 | 65 | 337 | 640 | 960 | 1629 | 3482 | 5750 |
| | 6 | 68 | 345 | 659 | 978 | 1652 | 3504 | 6054 |
| | 7 | 70 | 355 | 680 | 1007 | 1687 | 3591 | 6079 |
| | 8 | 71 | 363 | 685 | 1020 | 1745 | 3683 | 6230 |
| .50 | | | | | | | | |
| | 2 | 52 | 278 | 523 | 745 | 1413 | 2737 | 4685 |
| | 3 | 58 | 313 | 550 | 904 | 1401 | 3291 | 5114 |
| | 4 | 63 | 309 | 600 | 955 | 1517 | 3354 | 6011 |
| | 5 | 63 | 329 | 663 | 945 | 1580 | 3562 | 5991 |
| | 6 | 66 | 339 | 659 | 947 | 1711 | 3667 | 6255 |
| | 7 | 69 | 340 | 645 | 1037 | 1652 | 3418 | 6093 |
| | 8 | 73 | 361 | 682 | 1023 | 1798 | 3764 | 6382 |
| .80 | | | | | | | | |
| | 2 | 49 | 304 | 476 | 811 | 1316 | 2386 | 5202 |
| | 3 | 56 | 323 | 623 | 794 | 1679 | 3392 | 4654 |
| | 4 | 61 | 341 | 546 | 798 | 1318 | 3297 | 5381 |
| | 5 | 62 | 339 | 596 | 1026 | 1454 | 3533 | 5160 |
| | 6 | 64 | 316 | 610 | 953 | 1553 | 3881 | 5450 |
| | 7 | 70 | 377 | 669 | 972 | 1852 | 3545 | 7033 |
| | 8 | 68 | 324 | 634 | 1085 | 1856 | 3835 | 5545 |

## References

Cohen, J. Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press, 1969.

Glass, G. V. and Stanley, J. C. Statistical Methods in Education and Psychology. Englewood Cliffs: Prentice-Hall, 1970.

Hays, W. C. Statistics for the Social Sciences Holt, 1973.

Kirk, R.E. Experimental Design: Procedures for the Behavioral Sciences. Monterey: Brooks/Cole, 1984.

Meyer, P. Introductory Probability and Statistical Applications Reading: Addison-Wesley, 1965.

Pearson, E. S. and Hartley, H. O. Biometrika Tables For Statisticians. Cambridge,England: University Press, 1970.

Sheffe, H. The analysis of Variance. New York: Wiley, 1959.

Woodward J. A. and Overall, J. E. A computer Program for Calculating Power of the F-test. Educational and Psychological Measurement, 1977,36:165-167.